Trading Algorithms

Tyler Beason[1] **Sunil Wahal**[2]

We study primitive economic tradeoffs in trading between execution price, quantity traded, and time to trade. To do so, we employ proprietary data on institutional trading algorithms from a large provider of execution services. The firm is regarded as a leader in the field of electronic trading and has a large market-share. We study orders generated by four trading algorithms used by hundreds of institutions between 2012 and 2016. These algorithms are "standard" in the sense that almost all electronic brokers provide similar such trading algorithms.

The data provide extraordinary and hitherto unobservable insight into electronic trading. We observe a parent order originating from a buy-side desk/institution. Institutions are anonymous but uniquely identified. This allows us to control for, as well as examine, institution-specific effects. Unlike typical institutional trading data provided by SEI, the Plexus Group, and Ancerno, there is no ex post rolling up of trades to obtain parent orders. As such, our parent orders represent true ex ante demand.

For each parent, we observe the algorithm choice, side (buy, sell, or short-sale), ex ante order size, ex ante parent price limits imposed by the institution, ex ante parent volume limits imposed by the institution, and a host of other information. Once the parent order is transmitted by the Execution Management System (EMS) of the institution, it is received by the trading system of the electronic broker. The trading algorithm then breaks the parent into a large number of child submissions, some of which get filled. At the child order level, we observe order price type (market, limit, peg), start time, child size, limit price (if any), time in force, execution instruction, whether the order was hidden or displayed, submission venue, along with any fill statistics such as quantity filled, trade time, and price. For limit and pegged orders, we observe limit prices and type of peg (primary, midpoint, etc.). All data contain millisecond time-stamps for a variety of steps in the trading process such as submission time, receipt time at broker, execution time at matching engine, or cancellation.

The data are powerful for three reasons. First, because the algorithms that we study are not bespoke and are widely employed by all brokers, the results are generalizable to the market. Second, the data are "big". The sample consists of 2.3 million parent orders between 2012 and 2016 that represent $675 billion in aggregate demand across over 5,000 securities. These parent orders generate almost 300 million child submissions that represent $2.1 trillion in notional volume over this period. These child submissions ultimately result in $388 billion in traded volume. Since we do not receive market data from the broker, the data management and computational challenges grow by an order of magnitude as we add market data to the sample. This is especially the case because (a) the cross-section of securities is large (over 5,000 stocks), (b) the time series is long, and (c) the amount of market data generated for each child order is enormous. We have, so far, managed the data using a distributed cluster system but look forward to guidance XSEDE and other entities. Third, the data are unique. As described above,

[1] Email: Tyler.Beason@asu.edu, W.P. Carey School of Business, Arizona State University
[2] Email: Sunil.Wahal@asu.edu, W.P. Carey School of Business, Arizona State University

we observe the *entire* sequences of decisions made by trading systems (i.e. all child orders), not just the ex post executions. Most studies of institutional trading only examine realized trades because of data limitations.[3] It is particularly important to observe choices as well as outcomes because they are endogenously determined. Millisecond timestamps allow us to examine in detail in the microstructure of the market surrounding order submission and execution. Such investigations are simply impossible using conventional market data or other proprietary sources.

We have several questions of economic interest that we expect the data to be able to speak to. For some of these questions, analysis is well underway and we have a set of interesting results. We describe two of these below.

Tradeoffs in trading: Each algorithm encodes design features that exploit the market's microstructure to achieve its objective. We estimate partial correlations between child-level fill rates, time-to-trade, and trading costs, and a variety of hitherto unobservable microstructure choice variables. The set of microstructure choice variables that we examine are large. One example is whether an order is hidden or displayed. Hidden orders have lower priority, thereby reducing execution probability but may have lower price impact than displayed orders. The choice, therefore, represents a tradeoff between quantity and price. Since the choice is, by construction endogenous, we cannot recover the structural parameters. However, such partial correlations offer us a first look into this type of choice variable.

We can offer an abbreviated preview of some initial results. The vast majority of child orders are either limit or exchange-pegged (rather than synthetic pegged) orders. This suggests that traditional models of market versus limit order choice are less informative of the current market structure. Instead, models such as Li, Wang, and Ye (2019) in which buy-side algorithmic traders trade alongside faster HFTs are more germane to understanding price formation. These passive limit and pegged orders have positive price impact, consistent with Brogaard, Hendershott and Riordan (2019), but the impact depends crucially on the microstructure choice variables of the child order. The complexity of choices faced by market participants is often the subject of disagreement among regulators, academics and practitioners. Our results suggest that this complexity appears to reflect nuances in the price, quantity, and time tradeoffs that are primitive to trading rather than rent extraction.

Opportunity costs of not trading: The opportunity cost of not trading has been an object of interest since Perold (1988). Unfortunately, it has never been adequately measured because true ex ante demand was unobservable. Parent orders are either unfilled or partially filled in a meaningful fraction of our sample. This allows us to estimate the opportunity cost of not trading. Our approach is based on simulating counterfactuals, so our estimates are in fact bounds on the opportunity costs of trading. Nonetheless, they represent the first estimates of opportunity costs that have been elusive. Since we also observe revisions to parent orders (and their limits), the data also allows us to characterize tradeoffs made by humans as opposed to algorithms.

[3] The earliest studies employed data from SEI (Chan and Lakonishok (1995, 1997)). A subsequent set of papers used data from the Plexus Group (Keim and Madhavan (1995, 1997), Conrad, Johnson and Wahal (2001, 2003). A most recent set of papers employ data from Ancerno (see extensive list of citation in the survey paper by Hu, Jo, Wang, Xie (2018)). Finally, Frazzini, Israel, and Moskowitz (2018) employ data from a single buy-side institution.